The Thermodynamic Effects of Proline Residues in a Soluble Antiparallel Coiled Coil Model Protein

C500 final Report

Author: Vincent M. Waldman
PI: Dr. Martha Oakley PhD

Spring 2006

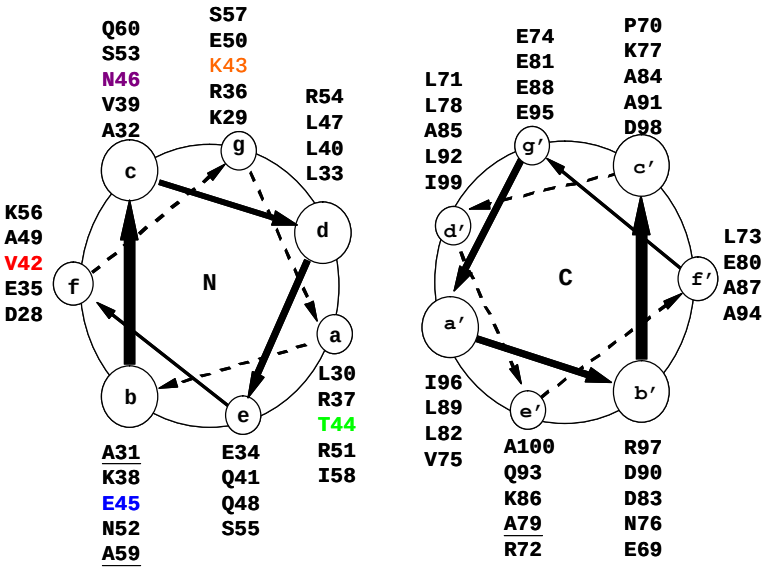
The goal of this project is to probe the thermodynamic effects of inserting a proline residue between positions of the helical register of a soluble antiparallel coiled coil. It has been demonstrated that proline residues are highly destablilizing within the context of isolated α-helices (*1*, *2*). Coiled coils are the juxtaposition of two or more α-helices. In part because of this relationship proline residues are typically regarded as coiled coil breakers. The amino acid chains that compose coiled coils are generally characterized by amphiphatic heptad repeats. The convention is to label each residue within the heptad **a**-**g** with positions **a** and **d** as hydrophobic and found at the helical interface.(*3*) For reasons discussed below it is our hypothesis that insertions of proline residues between specific register positions in the helix may be tolerated.

The thermodynamic penalty for including an α-helix in a coiled coil is between ~3-3.8 kcal mol^{-1} depending on which amino acid it is compared to (*1*, *2*). This is a result of the unique union of the backbone and side chain of a proline residue into a pyrolidine ring. The consequences of this structure include: 1) A proline residue has no amide proton available for participation in the extended hydrogen bonding network of an α-helical backbone. 2) Proline residues have little thermodynamic advantage to exist in the typical *trans* conformation as opposed to the *cis*. And 3) a proline residue influences the torsion angle of the residue preceding it generally causing the local conformation to be more extended than is typical in an α-helix (*4*, *5*). There is one exception to this; a proline residue can be stabilizing if found within the first four amino-terminal positions of an α-helix. Residues within a α-helix have an average torsion angle of $\phi = -57°$. The 5-membered ring of a proline residue can hold a similar torsion angle of $\phi = -60°$stably. Since the first four residues cannot act as hydrogen bond donors in an *i*+4 helix these positions generally allow the presence of a proline residue without much of a penalty. The first position is especially

accommodating to a proline residue as it does not have any preceding helical residues that would be affected in its presence (*6*).

Because their sequential periodicity makes them amenable to bioinformatics techniques and because of their prevalence in nature the prediction of coiled-coil structures has been an actively pursued topic. Indeed, coiled coil prediction algorithms exist that score amino acid side chains for coiled coil propensities. The representative samples used to develop such algorithms show that proline residues have a low propensity for existing in stable coiled coils (*7*, *8*). These algorithms leave much to interpretation however in that they are designed to seek out only canonical heptad repeats. If a sequence does not fit into such a repeat the algorithm will assign breaks in the coiled coil to correct the alignment. Proline residues most often are characterized within these breaks in structure. It is the author's hypothesis that many coiled coils may include proline residues without significant losses in stability. Such inclusion may lead to local perturbations in sequential periodicity, making them more difficult to predict than canonical stretches of coiled-coil structure, but may leave the global structure of the coiled coil intact.

a)



b)

Helical position	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*a*
Protein Code	R_{37}	K_{38}	V_{39}	L_{40}	Q_{41}	V_{42}	K_{43}	T_{44}	E_{45}	N_{46}	L_{47}	Q_{48}	A_{49}	E_{50}	R_{51}
DNA nontemplate	cgt	aaa	gta	ttg	cag	gtc	aaa	acg	gaa	aac	ctg	caa	gcg	gag	cgt

Figure 12: a) A helical wheel diagram of SRS-82. The colored residues V_{42}, K_{43}, T_{44}, E_{45} and N_{46} are the positions between which proline residues will be inserted as guest residues. The residues A31, A59 and A79 were mutated from glycine residues. b) An alignment of the helical positions of the highlighted residues with the protein and the non-template strand DNA sequences of SRS-82.

There is structural evidence to support that proline residues may exist in a slightly perturbed yet stable coiled coil. A crystal structure of GreA transcript cleavage factor from *E. coli* shows a centrally located proline residue in an antiparallel coiled coil domain in the **b** position of the helix. Flanking this proline residue are conserved salt bridges and an amino acid insertion in the heptad repeat between the preceding **f** and **g** positions. These features *may* make up for the thermodynamic penalty for the inclusion of the proline residue, though the author is unaware of any studies that have directly measured this. (*9*)

There is also evidence that proline residues can be incorporated readily into membrane bound helices (*6*, *10*). The hydrophobic environment of membranes is more amenable to the

extended conformation of the residue preceding a helical proline residue. (*4*, *5*) Also the residue in the *i*+4 position before the proline may be more relaxed than normal, favoring a hydrophobic environment, due to the imino group of the proline backbone that cannot act as a hydrogen bond donor.

Of interest to our study is the one amino acid insertion in the helical register of GreA as described above. This insertion allows the proline-including helix to relax its wind and pucker so that the overall structure of the coiled coil can be maintained (*9*). The design of our study mimics the primary structure of GreA to a first approximation. In this study proline residues will be inserted as heptad interrupting residues between the intra-heptad positions **a** and **b**, **b** and **c** and **f** and **g** and the inter-heptad position **g-a**. The other possible insertions between the positions **c** and **d**, **d** and **e** and **e** and **f** are roughly symmetrical to positions **a** and **b**, **f** and **g** and **g** and **a** respectively so these constructs were not made in the course of this study.

The protein scaffold that has been chosen to host these proline insertions is SRS-82. This protein is an 82-amino acid long construct of the coiled coil domain from seryl tRNA synthetase from *E. coli* that autonomously folds into an antiparallel dimeric coiled coil composed of one subunit. This protein expresses well under standard conditions and has a reasonable $\Delta G^{\circ}_{\text{folding}}$ of -3.6 kcal/mol for such a model (*11*). Since the addition of a proline residue is likely to be destabilizing the reasonably low free energy of folding in this construct is attractive for this study.

To increase the likelihood of experimental success the thermodynamic stability of SRS-82 was enhanced prior to inserting proline residues so that the stability of these insertions could more easily be assessed. This was done through the mutation of three intrahelical glycine residues to alanine residues. Glycine residues are typically destabilizing within a helix because of their

relatively high degree of conformational freedom and because their size decreases the amount of helical van Der Waal contacts and increases the solvent exposed surface area of the helix (*12*).

The overall goal of this project is to measure the change in thermodynamic stability of our antiparallel coiled coil model as a function of inserted proline residue position within a heptad. This putative penalty would then be assessed as:

$$\Delta\Delta G_{Proline\ Insertion} = \Delta G_{\ Unfold\ Proline\ Mutant} - \Delta G_{Unfold\ Wild\ Type}$$

The work presented in this study will show that our mutants are currently unsuitable for such an assessment but that it may still be possible to perform with some minor changes to our protocol.

Materials and Methods:

Protein sequences

The proteins designed for and studied in this work are sequentially similar. Table 1 shows these sequences and gives the logic behind how they are named.

Table 5: Protein sequences and names of the proteins used in this study. In table below the (^) denote thrombin cleavage sites, after thrombin digestion the protein names drop the H. The positions marked (A) represent where the intrahelical glycine residues from SRS-82 (*11*) were mutated to alanine. The residues represented with the (P) denote where proline residues were inserted into the helix.

HS3A (His-SRS 3A):
MGSSHHHHHHSSGLVPR^GSHMLDVDKLAALEERRKVLQVKTENLQAERNSRS KSIAQAKARGEDIEPLRLEVNKLAEELDAAKAELDALQAEIRDIALTIPW
HS3A-PFG (His-SRS 3A with a proline insert between **f** and **g** positions):
MGSSHHHHHHSSGLVPR^GSHMLDVDKLAALEERRKVLQVPKTENLQAERNSR SKSIAQAKARGEDIEPLRLEVNKLAEELDAAKAELDALQAEIRDIALTIPW
HS3A-PGA (His-SRS 3A with a proline insert between **g** and **a** positions):
MGSSHHHHHHSSGLVPR^GSHMLDVDKLAALEERRKVLQVKPTENLQAERNSR SKSIAQAKARGEDIEPLRLEVNKLAEELDAAKAELDALQAEIRDIALTIPW
HS3A-PAB (His-SRS 3A with a proline insert between **a** and **b** positions):
MGSSHHHHHHSSGLVPR^GSHMLDVDKLAALEERRKVLQVKTPENLQAERNSR SKSIAQAKARGEDIEPLRLEVNKLAEELDAAKAELDALQAEIRDIALTIPW
HS3A-PBC (His-SRS 3A with a proline insert between **b** and **c** positions):
MGSSHHHHHHSSGLVPRG^SHMLDVDKLAALEERRKVLQVKTEPNLQAERNSR SKSIAQAKARGEDIEPLRLEVNKLAEELDAAKAELDALQAEIRDIALTIPW

Plasmid Construction:

The first plasmid made for this study was a double alanine (2A) mutant constructed using site directed mutagenesis (SDM) as described below derived from the SRS-82 gene in the pAED4 vector (a gift from Martha Oakley and Peter S. Kim) (*13*). First the mutation, G59A, was made using the primers cgatcgaaatccatt*GCG*caggcgaaag and its complement, where the bases in uppercase italicized letters represent the site of mutation and the lowercase letters anneal to the

template. This plasmid was then used as a template to perform the G79A mutation with the primers ggaagtgaacaaactg*GCG*gaagagctggatg and its complement. The resulting clone, pSRS-82 double alanine (pS2A), as each of the clones in this study, was handled using standard molecular biological techniques and was verified with sequencing through the Indiana Molecular Biology Institute (*14*). All primers in this study were purchased through Sigma Genosys.

The plasmid pS2A was then used as a template for each of the four proline insertion mutations performed through SDM. The following primers and their compliments were used to insert a proline residue between the following positions in pS2A (residue numbers are derived from SRS-82): between residues V42 (**f**) and K43 (**g**) gtaaagtattgcaggtc*CCG*aaaacggaaaacctgcaag, between residues K43 (**g**) and T44 (**a**) gtaaagtattgcaggtcaaa*CCG*acggaaaacctgcaag, between residues T44 (**a**) and E45 (**b**) gtattgcaggtcaaaacg*CCG*gaaaacctgcaag and between residues E45 (**b**) and N46 (**c**) ggtcaaaacggaa*CCG*aacctgcaagcg.

After initial attempts to express and purify these mutants it was decided that these constructs were inadequate for this study. The double alanine mutants expressed poorly and it was difficult to separate this low yield from contaminants (data not shown). To improve these mutants their genes were cloned into pET 28 (Novagen) to put a 6xHis tag on their N-termini and the mutation G31A was performed on each yielding the triple alanine mutants shown in table 1. These point mutations were performed via a series of amplification PCR reactions (see below for details) resulting in mutated amplified DNA products between the flanking T7 promoter primer and the T7 terminator primer (Novagen). The primers used to perform the G31A mutations were ctggatgtagataagctg*GCG*gctcttgaagagcg and its complement. These products were then cloned into pET 28 (Novagen) between the restriction enzyme sites Nde I and Eco RI (See below for

general cloning procedure). The addition of the 6xHis tag and the exploitation of the T7 expression system of the pET 28 vector allowed the expression of an abundant amount of protein that could subsequently be purified with a nickel affinity column.

Site Directed Mutagenesis

The SDM polymerase chain reactions (PCR) were all prepared in 50 µL of total volume. Each reaction contained 2.5 Units of *PfuUltra* HF DNA polymerase (Stratagene), 1x *PfuUltra* HF DNA polymerase buffer (Stratagene), 125 µM each dNTP (New England Biolabs), 500 nM each appropriate primer, ~1-5 µg/mL DNA template and ddH_2O to volume. The PCR was performed in 24 cycles of 1 min. at 95°C, 1 min. at 65°C and 4 mins. at 72°C.

Amplification PCR

All amplification PCR reactions were prepared in 100 µL of total volume. Each reaction contained 2 units VENT DNA Polymerase (New England Biolabs), 1x Thermopol buffer (New England Biolabs), 2 mM $MgSO_4$, 2 µM *each* primer used, ~1-5 µg/mL DNA template, 500 µM each dNTP and ddH_2O to volume. The PCR used to perform the G31A mutations was carried out in three steps: 1) the DNA was amplified between the T7 Promotor primer and the mutagenesis primer that annealed to the non-template strand of the gene. 2) In a likewise fashion the DNA was amplified between the T7 terminator primer and the mutagenesis primer that annealed to the template strand of the gene. 3) The amplified products in steps 1) and 2) were agarose gel purified (see cloning procedure for more detailed description) then combined and used as templates/primers (at the primer concentration) for a third round of amplification between the T7 promoter and T7 terminator primers. All PCR reactions of this nature were performed in 30 cycles of 1 min. at 95°C, 1 min. at 65°C and 1 min. at 72°C.

Double Digest and Cloning

The amplified DNA products were isolated via agarose gel electrophoresis using standard techniques (*14*) and were purified using a QiaQuick gel extraction kit (Qiagen) following the protocol provided by the manufacturer. During this procedure the DNA was eluted from the spin columns used with the kit with ddH_2O in one half the volume of the volume of PCR reaction solution loaded into the gel. A volume of 10 µL of each purified DNA product was then double digested with Nde I and Eco RI in 20 µL reactions following the manufacturers suggested protocol. A volume of 10µl of a miniprep stock of pET 28 was likewise digested concurrently. The restriction digested DNA was isolated on an agarose gel and gel purified as before. The digested products were then ligated together in ratio of gene insert to plasmid backbone of 3 µL:1 µL in a 10µL total reaction volume. The ligation reactions used 400 units of T4 DNA Ligase in 1x T4 DNA ligase buffer and were allowed to progress for 1 hour at room temperature. The resulting DNA plasmids were then handled using standard molecular biological techniques (*14*) and were verified with sequencing. The restriction enzymes, the ligase and their buffers were purchased from New England Biolabs.

Expression and Purification

All proteins in this study were expressed and purified in a similar manner. First the sequence confirmed plasmids were transformed into BL21(DE3) cells (Novagen) following manufacturer's protocol. These cells were then plated onto Luria-Bertani (LB) broth supplemented with agar and 25 µg/mL kanamycin (kn). These were allowed to incubate overnight at 37°C. Individual colonies were used to inoculate 4 mL LB Kn overnight cultures. After ~16 hours in a 37°C incubator-shaker these cultures were used to inoculate 700 ml LB kn cultures in 2

L baffled flasks. The cells were allowed to reach an OD_{590} of 0.6 before induction with 400 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). After induction the cultures were allowed to grow for an additional ~5 hours before harvesting the cell pellet with centrifugation. The cell pellets were then frozen at −70°C before further processing.

For purification the cells were first thawed and then resuspended in ~40 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0) supplemented with ~1 mg/mL lysozyme (Sigma-Aldrich). The cells were then lysed with a fluidizer. To The lysate ~70 μM phenylmethylsulfonyl fluoride (PMSF) was immediately added to guard against serine protease degradation. The soluble fraction was then isolated through centrifugation and this was applied to a 4 mL Ni-NTA Agarose column (Qiagen). The flow through was collected and the column was washed with ~10 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0). The proteins were then typically eluted into 5 mL fractions with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0).

The mutants HS3A-PFG, HS3A-PGA and HS3A-PBC were further purified as follows. The proline insertion mutants tended to aggregate when buffer conditions deviated from pH 8.0. Because of this it was necessary to solubilize these mutants with 2 M guanidinium hydrochloride. The proteins were then purified with reverse phase HPLC, using a Vydac C_{18} column and a linear water/acetonitrile gradient containing 0.1% trifluoroacetic acid. After HPLC excess acetonitrile was removed with a rotovap and the protein solutions were lyophilized to powder. About 5 mg of each of these proteins were then resuspended in 20 mL of 20 mM tris buffer at pH 8.0 and were digested with 5 units of thrombin (Sigma) overnight at room temperature to remove the 6xHis tag. The proteins were again purified with reverse phase HPLC, rotovapped and lyophilized to powder. The molecular masses of these proteins were verified by matrix-assisted laser

desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry in the Indiana University Chemistry Department Mass Spectrometry Facility. All molecular weights were within 0.1% of the expected masses.

The mutant HS3A-PAB was not completely purified during the course of this study. It should be noted that this protein expressed successfully but was lost while optimizing the procedures for handling the proline insertion mutants. It should be possible however to purify this protein following the procedures above.

The HS3A protein was immediately dialyzed into 10 mM PBS buffer pH 8.5 following elution from the Ni-NTA Agarose column. Once equilibrated the protein was digested with 5 units of thrombin (Sigma) overnight at room temperature. The cleaved protein was separated from thrombin via a 400 mL Sephadex S50 size exclusion chromatography column. Fractions containing S3A (gel not shown) were pooled, dialyzed into ddH_2O and were lyophilized to powder. The molecular mass of the S3A mutant was then verified with MALDI-TOF mass spectrometry as before.

Characterization with Circular Dichroism

The proteins S3A, S3A-PFG, S3A-PGA and S3A-PBC were resuspended in phosphate buffered saline (PBS, 10 mM phosphate, 150 mM NaCl, pH 8.0) and the concentration of each protein stock was determined in 3 M guanidinium hydrochloride measuring the absorbance at $A_{260.}$ The extinction coefficients for the proteins were calculated to be 5690 $M^{-1}cm^{-1}$ using the method outlined by Gasteiger et. al.(*15*) Each of the protein stocks were diluted to roughly 10 μM with buffer for CD experiments.

The CD experiments were performed on a Jasco J-715 spectropolarimeter. The spectrum for each peptide was corrected for non-peptide absorbance by subtracting out the spectrum of PBS resuspension buffer only. Wavelengths scans were performed on the samples at 2°C from 200 nm to 270 nm in 0.5 nm increments. These scans were averaged 10 times and were performed in a 0.5 cm cell. Thermal stabilities of the proteins were assessed by measuring the ellipticity of the sample at 222 nm as a function of temperature. The temperature was increased from 2°C to 90°C in 2°C increments. After the instrument reached a specific temperature the sample was allowed to equilibrate for 90 seconds before measurements were made as an average of 5 scans. The sample cuvette was equipped with a stir bar and lid for these measurements.

***Reesults*:**

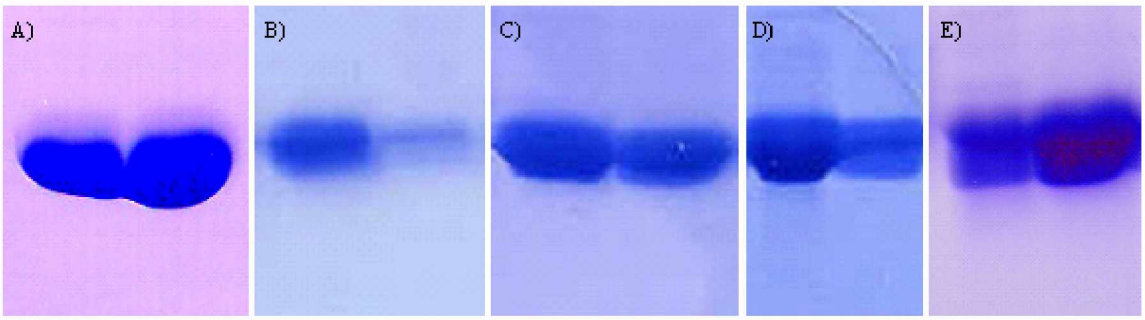


Figure 13: A comparison of expression levels and Ni-NTA Agarose column recovery of A) HS3A, B) HS3A-PFG, C) HS3A-PGA, D) HS3A-PAB and E) HS3A-PBC. The sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) coomassie blue stained gels shown correspond to samples of the first two Ni-NTA Agarose column elution fractions of each purified proteins. Each of these proteins was purified through this step under non-denaturing conditions. All of the samples were collected in a similar fashion but it should be noted that this figure is suitable for *qualitative* comparisons only.

During expression and purification each of the HS3A mutants expressed robustly. In general, the majority of each protein eluted from the Ni-NTA Agarose column in a volume of 10 ml of elution buffer; this corresponds to the first two elution fractions (see figure 2). The yields of

the proteins before the thrombin digest were each estimated at over 50 mg per 1.4 L of expression media (results not shown).

Using circular dichroism Oakley and Kim showed that their construct, SRS-82, was about 85% helical at 0°C and had a thermal unfolding transition at ~50°C. (*11*) Using the value $[\Theta]_{222}$ = –30266 deg cm^2 dmol^{-1} for the protein S3A (see figure 3) the helical content of this protein was determined to be 79% at 2°C using the method of Chen. *et. al.* (*16*) This compares well with helical content of SRS-82 when considering these proteins were analyzed under slightly different conditions. In addition, S3A has an additional three residues (Gly-Ser-His) bound to its N-terminus that are remnants from the thrombin cleaved 6xHis tag. These residues may not be in helical conformations which would lead to a slight decrease in the percent helicity of S3A relative to SRS-82. The thermal unfolding transition of S3A is estimated to be between 57°C and 61°C (see figure 4 and inset of figure 5). Therefore we succeeded in increasing the thermal stability of SRS-82 through the G31A, G59A and G79A mutations to make the construct S3A.

The wavelength scans of the proline insertion mutants (See figure 3) indicate that each of these proteins have a significant decrease in helical content at 2°C relative to S3A. The molar ellipticities at 222 nm of S3A-PFG, S3A-PBC and S3A-PGA were determined to be –9738, –14718 and –11994 deg cm^2 dmol^{-1} respectively. These values were used to determine the helical content of these mutants to be 25%, 38% and 31% in the same order.*(16)* As is the case with S3A these proteins had an additional three residues that were remnants from the thrombin cleaved 6xHis tag, therefore these percentages are likely slightly less than they should be relative to SRS-82. In addition to the marked drop of percent helical content the shapes of these scans indicate that these mutants have both helical and unordered character. The unordered character of these

proteins is revealed with their global minima between 204 and 206 nm having much greater magnitudes than their local minima at 222 nm. *(16)*

The thermal scans show that the proline insertion mutants unfold at much lower temperatures than S3A (See figure 4). A first derivative plot of figure 4 was constructed (see figure 5) to estimate these thermal transitions more accurately. In addition to melting at lower temperatures the thermal transitions of the proline mutants are much less distinct than that of S3A. Indeed it is difficult to determine if these proteins existed in a unique folded structure before reaching their estimated melting temperatures. These thermal transitions are estimated to be between 5°C and 25°C for S3A-PBC, 20°C and 40°C for S3A-PGA and 15°C and 35°C for S3A-PFG. Taken together figures 3-5 suggest that the proline mutants are somewhat unfolded at 2°C though further testing would be required to verify this.

To reiterate, figure 2 shows that the 6xHis tagged proline insertion mutants survived in abundant quantities through nickel ion chromatography under non-denaturing conditions. Because of this it was speculated that the 6xHis tag may impart some stability to the mutants. To check this, CD spectra were collected for the protein HS3A-PGA to compare against the spectra for S3A-PGA. The molar ellipticity values for HS3A-PGA shown in figure 3 were calculated taking into consideration the length of the entire protein. As shown the molar ellipticity of HS3A-PGA at 222 nm is -9728 deg cm^2 $dmol^{-1}$, this corresponds to a helical content of 25%. One can normalize this calculation against S3A-PGA if it is assumed that the contribution of the 6xHis tag to the helicity of the protein is negligible. To do this it is assumed that the helical content of HS3A-PGA is derived from the same number of residues as in S3A-PGA. If molar ellipiticity and the percent helicity are calculated in this way the molar ellipticity of HS3A-PGA at 222 nm decreases to -11674 deg cm^2 $dmol^{-1}$ corresponding to an increase in helicity to 30%. This normalized value

compares closely to the 31% helical content of S3A-PGA suggesting there is little deviance in helical content between the putative coiled-coil domains of these proteins. Additionally, if the molar ellipticity at 222 nm of HS3A-PGA is normalized to the length of S3A-PGA the thermal scan of this protein compares very closely to that of S3A-PGA (see figure 4). The slight differences in shape of the thermal transition of HS3A-PGA, however suggests this protein thermally denatures between 5°C and 30°C. This estimation is a lower temperature range than S3A-PGA. (see figures 4 and 5). Therefore, it is unlikely that the 6xHis tag bolsters the stability of the S3A-PGA mutant.

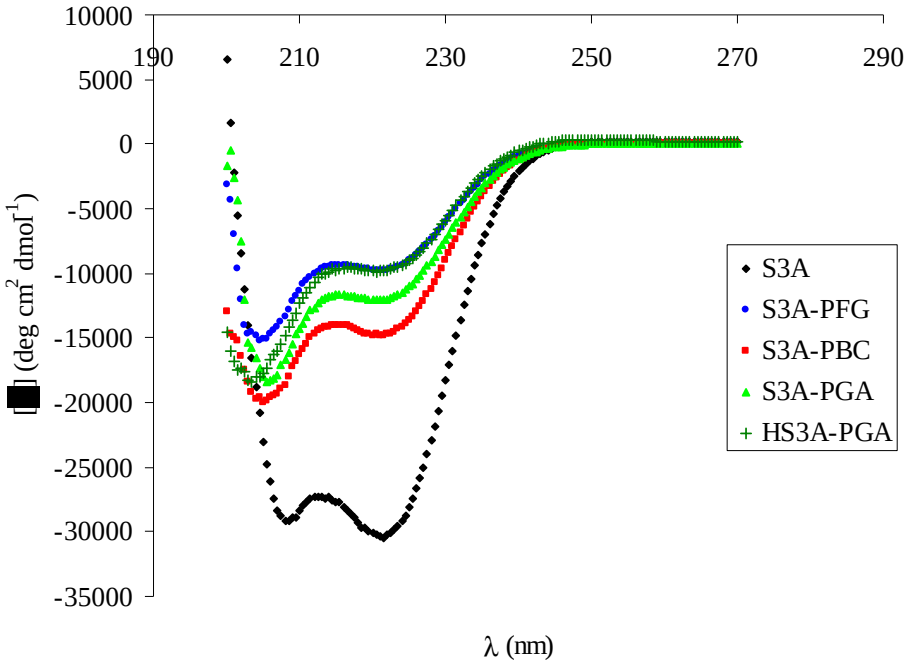


Figure 14: Wavelength scans from 200 nm to 270 nm of S3A, the S3A proline insertion mutants and HS3A-PGA for comparison with S3A-PGA.

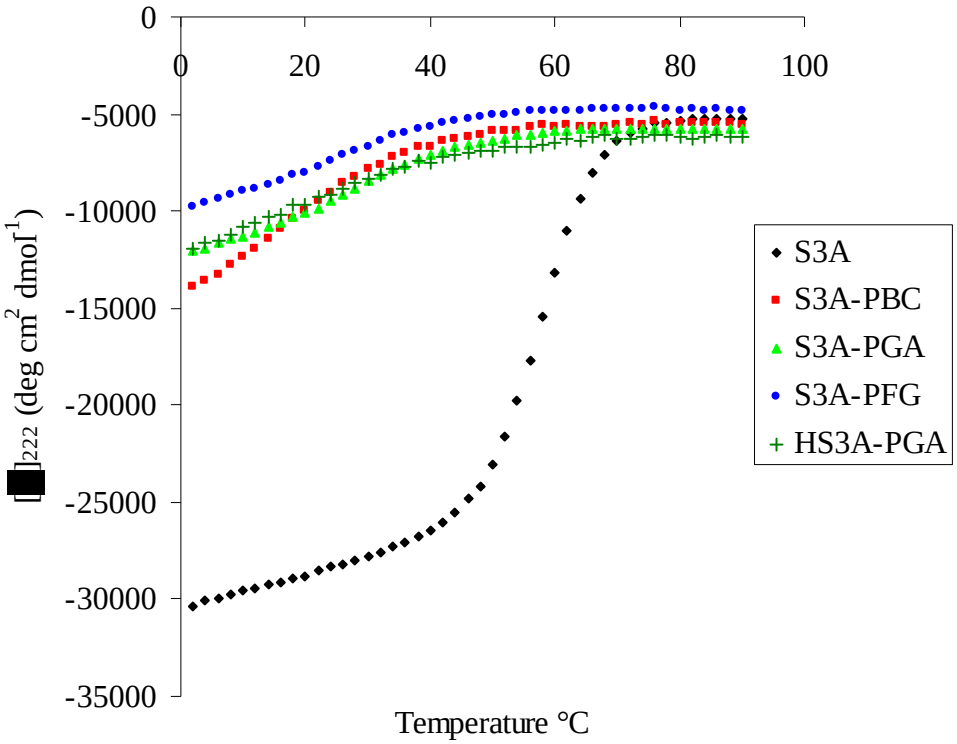


Figure 15: Thermal scans of S3A and the S3A proline insertion mutants from 2°C to 90°C at two degree intervals at 222nm. A scan of the protein HS3A-PGA is included for comparison against S3A-PGA. The molar ellipticity of this protein was normalized to the length of S3A-PGA with the goal of making the comparison more meaningful.

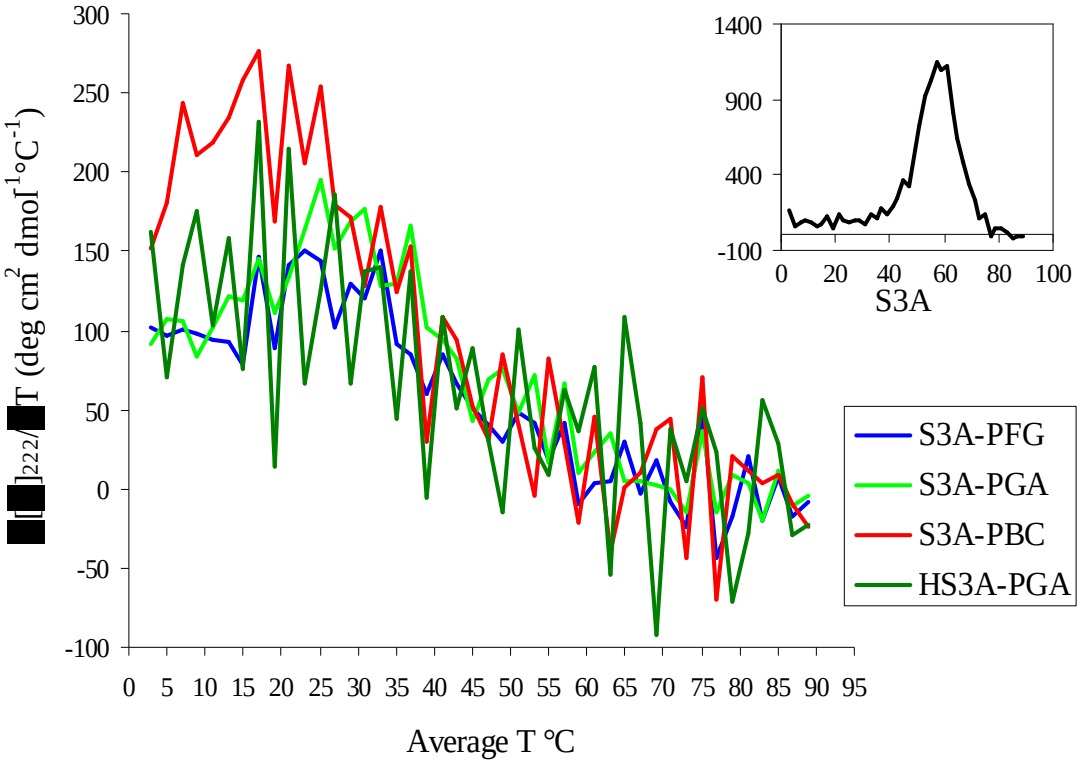


Figure 16: A first derivative plot of the thermal melts shown in figure 4. This plot was used to estimate the temperature ranges of each of the mutants in this study. The first derivative plot of S3A is shown as an inset because the scale of this plot makes it difficult to evaluate the transitions of the proline insertion mutants when plotted together. The axes in the inset have the same units as those the main plot.

Conclusions and future work:

We were successful in enhancing the thermal stability of SRS-82. Through the mutation of three interhelical glycine residues to alanine residues we increased the thermal stability from a T_m of ~50°C in SRS-82 to a range of 57°C-61°C with our S3A construct. This property may make this construct useful for studies.

It is clear from Figures 3-5 that the proline insertion mutants S3A-PFG, S3A-PGA and S3A-PBC, will be inadequate for the calculation of $\Delta\Delta G_{Proline\ Insertion}$. It is also reasonable to assume that S3A-PAB will also be too instable to make meaningful thermodynamic

measurements. On the basis of the CD results, it is puzzling that these proteins expressed well and were stable enough to survive purification through Ni-NTA Agarose chromatography (see figure 2). The presence of a significant amount of unordered signal in the spectra of these proteins and their relatively low melting temperatures suggest that the proteins are somewhat unfolded, though other experiments would be required to verify this. This would be especially true at the temperatures of expression and purification. Why these proteins seemed to be robust against proteolytic degradation then is up for speculation.

It was thought that the 6xHis tag may impart some stability to the proline insertion constructs. Indeed, initial attempts to express and purify the S2A predecessors to the HS3A proteins from the pAED4 vector proved difficult (results not shown). The addition of the 6xHis tag was performed to ease purification not to increase the protein stability. Therefore when expression improved after the genes were cloned into the pET 28 hosts it was reasoned that this was a result of the vector change and the G31A mutations that were placed into the genes through the cloning procedure. To make sure that the 6xHis tag was not a stabilizing appendage the protein HS3A-PGA was scanned for comparison against S3A-PGA (Figure 3). This comparison shows that percent helical content of the structure deviates little with or without the 6xHis tag.

In order to achieve our goal of evaluating the thermodynamic effects of including a proline residue in a coiled coil a different approach will have to be taken. It is possible that the structural stability of a coiled coil upon inclusion of a proline residue is supremely sensitive to the shifts in the register made. In the coiled-coil domain of GreA the proline residue occupies a **b** position with a one residue break in the register of the helix occurring between the preceding **f** and **g** positions. This break is occupied by a valine residue. It is more likely, however, that this structure

is stabilized through a number of side chain interactions, including both salt bridges and hydrogen bonds, that flank the proline residue. The salt bridges or hydrogen bonds circa Pro 27 in GreA include Glu 17-Arg 61, Arg 25-Asn 54, Arg 26-Glu 59, Arg 26- amide oxygen of Ala 36 and Lys 22-Glu 66. (*9*)

Perhaps another route to study the role of a proline residue in this context would be to begin with a construct of the coiled-coil domain of GreA. The structure of GreA consists of two domains the N-terminal antiparallel coiled-coil domain and a C-terminal domain composed of an α-helix cradled in a 5-strand β-sheet. Each of these domains contributes to roughly half the 158 residues of the protein. (*9*) It would be interesting to see if the coiled coil of N-terminal domain that contains the proline insertion can fold autonomously. If it does the aforementioned salt bridges and hydrogen bonds could then be replaced with conservative mutations to assess their role in stability. Once a more complete understanding of how GreA incorporates a proline insertion is known a more rational approach can be taken to build constructs to monitor proline residues in coiled coils.

Acknowledgements

I'd like to thank a number of people who have contributed to this project. First I'd like to thank Zac Kaur who performed much of the labor in the cloning, expression and purification of the HS3A mutants. I'd also like to thank Brittney Greene and Sarah Teeter, two very intelligent undergrads, that I have had help from during the course of this project. Last I'd like to thank my PI Martha Oakley for allowing me to pursue and helping me develop this project and for her mentorship along the way.

References:

(1) O'Neil, K. T., and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*, 646-51.

(2) Blaber, M., Zhang, X. J., and Matthews, B. W. (1993) Structural basis of amino acid alpha helix propensity. *Science 260*, 1637-40.

(3) Woolfson, D. N. (2005) The design of coiled-coil structures and assemblies. *Adv Protein Chem 70*, 79-112.

(4) Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins 50*, 437-50.

(5) Schimmel, P. R., and Flory, P. J. (1968) Conformational energies and configurational statistics of copolypeptides containing L-proline. *J Mol Biol 34*, 105-20.

(6) Senes, A., Engel, D. E., and DeGrado, W. F. (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol 14*, 465-79.

(7) Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995) Predicting coiled coils by use of pairwise residue correlations. *Proc Natl Acad Sci U S A 92*, 8259-63.

(8) Lupas, A., Van Dyke, M., and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science 252*, 1162-4.

(9) Stebbins, C. E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A., and Darst, S. A. (1995) Crystal structure of the GreA transcript cleavage factor from Escherichia coli. *Nature 373*, 636-40.

(10) Li, S. C., Goto, N. K., Williams, K. A., and Deber, C. M. (1996) Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc Natl Acad Sci U S A 93*, 6676-81.

(11) Oakley, M. G., and Kim, P. S. (1997) Protein dissection of the antiparallel coiled coil from Escherichia coli seryl tRNA synthetase. *Biochemistry 36*, 2544-9.

(12) Fersht, A. (1999) *Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding*, W.H. Freeman, New York.

(13) Doering, D. S. (1992) pp 244 leaves.

(14) Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning : a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

(15) Walker, J. M. (2005) *The proteomics protocols handbook*, Humana Press, Totowa, N.J.

(16) Yee-Hsiung Chen, J. T. Y., and Kue Hung Chau. (1974) Determination of the helix and β form of proteins in aqueous solution by circular dichroism
Biochemistry 13, 3350 - 3359.